JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 114 (2000) 305–318

Convergence properties of Krylov subspace methods for singular linear systems with arbitrary index ☆

Yimin Wei[a,*], Hebing Wu[b]

[a] Department of Mathematics, and Laboratory of Mathematics for Nonlinear Sciences, Fudan University, Shanghai, 200433, People's Republic of China

[b] Institute of Mathematics, Fudan University, Shanghai 200433, People's Republic of China

Received 2 July 1998; received in revised form 29 April 1999

Abstract

Krylov subspace methods have been recently considered to solve singular linear systems $Ax=b$. In this paper, we derive the necessary and sufficient conditions guaranteeing that a Krylov subspace method converges to a vector $A^{\mathrm{D}}b+Px_0$, where A^{D} is the Drazin inverse of A and P is the projection $P=I-A^{\mathrm{D}}A$. Let k be the index of A. We further show that $A^{\mathrm{D}}b+Px_0$, $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, is a generalized least-squares solution of $Ax=b$ in $\mathscr{R}(A^k)+\mathscr{N}(A)$. Finally, we present the convergence bounds for the quasi-minimal residual algorithm (QMR) and transpose-free quasi-minimal residual algorithm (TFQMR). The index k of A in this paper can be arbitrary, which extends to the main results of Freund and Hochbruck (Numer. Linear Algebra Appl. 1 (1994) 403–420) that only considers the case $k=1$. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 15A09; 65F10; 65F20

Keywords: Singular system; Index; Drazin inverse; Jordan canonical form; Krylov subspace method; Generalized least-squares solution

1. Introduction

We consider the general singular linear systems [5,7,11,14]

$$Ax=b, \tag{1.1}$$

where $A\in\mathscr{C}^{n\times n}$ and $b\in\mathscr{C}^n$. To exclude trivialities, we always assume $n\geqslant 2$.

☆ This research is supported by National Nature Science Foundation of China and the Youth Science Foundation of the Universities of Shanghai in China; the second author is also supported by the State Major Key Project for the Basic Researches and the Doctoral Point Foundation of China.

* Corresponding author.

E-mail addresses: ymwei@fudan.edu.cn (Y. Wei), 960167@fudan.edu.cn (H. Wu)

PII: S0377-0427(99)00237-X

An iterative scheme for solving linear systems (1.1) is called a Krylov subspace method if it produces approximate solutions of the form

$$x_m \in x_0 + \mathscr{K}_m(A, r_0), \quad m = 1, 2, \ldots, \tag{1.2}$$

where x_0 is an arbitrary initial guess with the corresponding residual vector $r_0 = b - Ax_0$ and $\mathscr{K}_m(A, r_0)$ is the mth Krylov subspace defined by

$$\mathscr{K}_m(A, r_0) = \mathrm{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}. \tag{1.3}$$

There are two main steps in designing a Krylov subspace method. The first step is the construction of suitable vectors $v_1, v_2, \ldots, v_m$ that span $\mathscr{K}_m(A, r_0)$. Then, setting

$$V_m = [v_1, v_2, \cdots, v_m], \tag{1.4}$$

we parameterize the mth iteration (1.2) as follows:

$$x_m = x_0 + V_m z_m, \quad \text{where} \quad z_m \in \mathscr{C}^m. \tag{1.5}$$

Therefore, as a second step, it remains to specify the choice of z_m in (1.5). Various strategies in the above two steps lead to various Krylov subspace methods, such as the generalized minimal residual algorithm (GMRES) [20], biconjugate gradient algorithm (BICG) [19], QMR [12] and TFQMR [9], etc. We shall not be more specific about the details of a Krylov subspace method but assume that it successfully generates the iterations $\{x_m\}_{m \geqslant 0}$.

To solve the singular linear systems (1.1), many methods have been suggested. Among them are the matrix splitting methods [16,17] and semiiterative methods [6,7,13,21], for a survey, see [7]. Recently, much attention is drawn to the Krylov subspace methods. An important feature of the Krylov methods is that they involve only multiplications of vectors with matrices A and A^*. It is noted that GMRES [1,18] as well as QMR and TFQMR [11] can be used to solve singular homogeneous systems that arise in the Markov chain modeling. Also deflation-like modifications of GMRES based on the truncated singular value decomposition have been considered in [15] and references therein. Special results were obtained for these methods, (cf. [3,11]). However, these results are derived in the case of that where the index of A is one.

The purpose of this paper is to study more thoroughly on the use of the Krylov methods to solve the singular systems (1.1) when the index of A is greater than one. After a brief review of some preliminary results in Section 2, we discuss the convergence behavior of the general Krylov subspace method for the case $b \in \mathscr{R}(A^k)$ in Section 3 and the case $b \notin \mathscr{R}(A^k)$ in Section 4. Finally, we present the convergence bounds for the QMR and TFQMR algorithms in Section 5.

2. Notation and preliminaries

Throughout this paper, $\mathscr{C}^n$ denotes the n-dimensional complex space, $\mathscr{C}^{n\times n}$ denotes the $n \times n$ complex matrix and I denotes the identity matrix. For $x, y \in \mathscr{C}^n$, we denote (x, y) the inner product of x and y. Letting L and M be complementary subspaces of $\mathscr{C}^n$, we denote by $P_{L,M}$ the projection on L along M. For $A \in \mathscr{C}^{n\times n}$, $\mathscr{R}(A)$, $\mathscr{N}(A)$ denote the range space and the null space of A respectively, and $\sigma(A)$ denotes the set of all eigenvalues of A. As usual, A^* is the conjugate transpose of A. The notation $\|.\|_2$ stands for the Euclidean norm.

Now, we give the definitions for the index and the Drazin inverse.

Definition 2.1. *Ben-Israel and Greville* [2]. Let $A \in \mathscr{C}^{n\times n}$. The smallest nonnegative integer k such that

$$\operatorname{rank}(A^k) = \operatorname{rank}(A^{k+1}) \tag{2.1}$$

is called the index of A, and is denoted by $k = \operatorname{index}(A)$.

Definition 2.2. *Ben-Israel and Greville* [2]. Let $A \in \mathscr{C}^{n\times n}$ with $\operatorname{index}(A) = k$, and $X \in \mathscr{C}^{n\times n}$ be such that

$$AX = XA, \qquad A^{k+1}X = A^k, \qquad XAX = X. \tag{2.2}$$

Then X is called the Drazin inverse of A, and is denoted by A^{D}. In particular, when $\operatorname{index}(A) = 1$, the matrix X satisfying (2.2) is called the group inverse of A, and is denoted by $A^{\#}$. If A is nonsingular, it is easily seen that $\operatorname{index}(A) = 0$ and A^{-1} satisfies (2.2), i.e., $A^{\mathrm{D}} = A^{-1}$.

The Drazin inverse can be represented explicitly from the Jordan canonical form of A. Let *index* $(A) = k > 0$. Then there exists a nonsingular matrix S such that

$$A = S \begin{bmatrix} R & 0 \\ 0 & N \end{bmatrix} S^{-1}, \tag{2.3}$$

where R is a nonsingular upper bidiagonal matrix and N is nilpotent of index k, i.e., $N^k \equiv 0$, and

$$A^{\mathrm{D}} = S \begin{bmatrix} R^{-1} & 0 \\ 0 & 0 \end{bmatrix} S^{-1}. \tag{2.4}$$

The following two lemmata characterize the index of a matrix.

Lemma 2.1. Ben-Israel and Greville [2]. *Let* $A \in \mathscr{C}^{n\times n}$. *Then* $\operatorname{index}(A) = k$ *if and only if* $\mathscr{R}(A^k) \oplus \mathscr{N}(A^k) = \mathscr{C}^n$.

Lemma 2.2. Ben-Israel and Greville [2]. *Let* $A \in \mathscr{C}^{n\times n}$ *with* $\operatorname{index}(A) = k$, $l \geqslant k$. *Then,*
(a) $\mathscr{R}(A^{\mathrm{D}}) = \mathscr{R}(A^l)$, $\mathscr{N}(A^{\mathrm{D}}) = \mathscr{N}(A^l)$;
(b) $A^{\mathrm{D}}A = AA^{\mathrm{D}} = P_{\mathscr{R}(A^{\mathrm{D}}),\mathscr{N}(A^{\mathrm{D}})} = P_{\mathscr{R}(A^l),\mathscr{N}(A^l)}$.

For a singular matrix A with $\operatorname{index}(A) = k$, we assume that the minimal polynomial of A has the form

$$\chi(\lambda) = \lambda^k \prod_{i=1}^{q} (\lambda - \lambda_i)^{q_i}, \quad (\lambda_i \neq 0,\ \lambda_i \neq \lambda_j \ \textit{for}\ i \neq j) \tag{2.5}$$

Finally, we denote by

$$\mathscr{P}_m = \{ p_m(\lambda) = \gamma_0 + \gamma_1\lambda + \cdots + \gamma_m\lambda^m \,|\, \gamma_i \in \mathscr{C},\ i = 0, 1, \cdots, m \}, \tag{2.6}$$

the set of all the complex polynomials of degree at most m.

3. The case $b \in \mathscr{R}(A^k)$, k = index(A)

It is well known that the iterations (1.2) of any Krylov subspace method can be expressed in terms of polynomials:

$$x_m = x_0 + \varphi_{m-1}(A) r_0, \quad \text{where } \varphi_{m-1} \in \mathscr{P}_{m-1}. \tag{3.1}$$

The corresponding residual vector r_m is given by

$$r_m = b - Ax_m = (I - A\varphi_{m-1}(A)) r_0 = \psi_m(A) r_0, \tag{3.2}$$

where

$$\psi_m(\lambda) \equiv 1 - \lambda\varphi_{m-1}(\lambda). \tag{3.3}$$

Note that $\psi_m \in \mathscr{P}_m$ and $\psi_m(0) = 1$; any such polynomial is called an mth residual polynomial.

The Drazin inverse is an important tool for solving singular system (1.1). In general, the Drazin inverse is not an "equation solving" inverse. Campbell [4] showed that $A^{\mathrm{D}}b$ is a solution of system (1.1) if and only if $b \in \mathscr{R}(A^k)$, and pointed that $A^{\mathrm{D}}b$ is the unique solution in $\mathscr{R}(A^k)$. Furthermore, we have:

Lemma 3.1. *Let* $A \in \mathscr{C}^{n\times n}$ *with* index$(A) = k$ *and* $b \in \mathscr{R}(A^k)$. *Then the general solution of the system* (1.1) *is of the form*

$$x = A^{\mathrm{D}}b + (I - A^{\mathrm{D}}A)z, \forall z \in \mathscr{R}(A^{k-1}) + \mathscr{N}(A). \tag{3.4}$$

In particular, the unique minimal $\|.\|_S$ norm solution of system (1.1) is presented by

$$x_{\mathrm{opt}} = A^{\mathrm{D}}b, \tag{3.5}$$

where S is the nonsingular matrix in (2.3) and the norm $\|.\|_S$ is defined by $\|x\|_S = \|S^{-1}x\|_2$ for any $x \in \mathscr{C}^n$.

Proof. The first statement is obvious. Now, we only prove the second statement. Since

$$\begin{aligned}
\|x\|_S^2 &= \|S^{-1}[A^{\mathrm{D}}b + (I - A^{\mathrm{D}}A)z]\|_2^2 \\
&= \|S^{-1}A^{\mathrm{D}}b\|_2^2 + (S^{-1}A^{\mathrm{D}}b, S^{-1}(I - A^{\mathrm{D}}A)z) \\
&\quad +(S^{-1}(I - A^{\mathrm{D}}A)z, S^{-1}A^{\mathrm{D}}b) + \|S^{-1}(I - A^{\mathrm{D}}A)z\|_2^2 \\
&= \|A^{\mathrm{D}}b\|_S^2 + (S^{-1}A^{\mathrm{D}}b, S^{-1}(I - A^{\mathrm{D}}A)z) \\
&\quad +(S^{-1}(I - A^{\mathrm{D}}A)z, S^{-1}A^{\mathrm{D}}b) + \|(I - A^{\mathrm{D}}A)z\|_S^2.
\end{aligned}$$

From (2.3) and (2.4), we have

$$(S^{-1}A^{\mathrm{D}}b, S^{-1}(I - A^{\mathrm{D}}A)z) = (S^{-1}(I - A^{\mathrm{D}}A)z, S^{-1}A^{\mathrm{D}}b) = 0.$$

Thus,

$$\|x\|_S^2 = \|A^{\mathrm{D}}b\|_S^2 + \|(I - A^{\mathrm{D}}A)z\|_S^2 \geqslant \|A^{\mathrm{D}}b\|_S^2 = \|x_{\mathrm{opt}}\|_S^2.$$

The proof is complete. □

Naturally, it should be questioned that under what conditions the iterations $\{x_m\}_{m\geqslant 0}$ generated by a Krylov subspace method converge to a solution of system (1.1).

Theorem 3.2. *Let $A\in\mathscr{C}^{n\times n}$ with index $(A)=k$. Let $b\in\mathscr{R}(A^k)$ and the minimal polynomial of A be given by* (2.5). *Then the following four statements are equivalent*:

(i) *The Krylov iterative sequence $\{x_m\}_{m\geqslant 0}$ converges, for any $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, to a solution of the system* (1.1);

(ii) $\lim_{m\to\infty}\psi_m(A)v=0$ *for every* $v\in\mathscr{R}(A^k)$;

(iii) $\lim_{m\to\infty}\varphi_{m-1}(A)v=A^{\mathrm{D}}v$ *for each* $v\in\mathscr{R}(A^k)$;

(iv) $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$.

If one — and thus all of these conditions — are fulfilled, then

$$\lim_{m\to\infty} x_m=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0. \tag{3.6}$$

Proof. Note that, for all $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, there holds $r_0=b-Ax_0\in\mathscr{R}(A^k)$ because $b\in\mathscr{R}(A^k)$.

(i)⇔(ii): In view of (3.2), (i) is equivalent to $\lim_{m\to\infty}\psi_m(A)r_0=0$, that is, $\lim_{m\to\infty}\psi_m(A)v=0$ for every $v\in\mathscr{R}(A^k)$.

(ii)⇔(iii): With (3.3) and Lemma 2.2, we have for every $v\in\mathscr{R}(A^k)$,

$$\psi_m(A)v=(I-A\varphi_{m-1}(A))v=A(A^{\mathrm{D}}v-\varphi_{m-1}(A)v). \tag{3.7}$$

It easily follows from (3.7) that $\lim_{m\to\infty}\psi_m(A)v=0$ for every $v\in\mathscr{R}(A^k)$ is equivalent to

$$\lim_{m\to\infty}(A^{\mathrm{D}}v-\varphi_{m-1}(A)v)\in\mathscr{N}(A)\subseteq\mathscr{N}(A^k),\quad \forall v\in\mathscr{R}(A^k). \tag{3.8}$$

Since $A^{\mathrm{D}}v-\varphi_{m-1}(A)v\in\mathscr{R}(A^k)$ by Lemma 2.2, (ii) is therefore equivalent to (cf. Lemma 2.1)

$$\lim_{m\to\infty}\varphi_{m-1}(A)v=A^{\mathrm{D}}v\quad \text{for every } v\in\mathscr{R}(A^k). \tag{3.9}$$

(ii)⇔(iv): It is easy to see that $\lim_{m\to\infty}\psi_m(A)v=0, \forall v\in\mathscr{R}(A^k)$ is equivalent to $\lim_{m\to\infty}AA^{\mathrm{D}}\psi_m(A)=0$. In view of the Jordan canonical form of A (2.3), (ii) is then equivalent to

$$\lim_{m\to\infty}\psi_m(R)=0, \tag{3.10}$$

i.e., $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$.

The limit (3.6) follows directly from (3.1) and (iii). □

Lemma 3.1 and Theorem 3.2 imply that, under the conditions of Theorem 3.2, every solution of system (1.1) can be approximated by a Krylov subspace method only if the initial guess x_0 is appropriately chosen.

If $x_0\notin\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, the vector $A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ is not a solution of the system (1.1). However, we have

Theorem 3.3. *Let $A\in\mathscr{C}^{n\times n}$ with* index $(A)=k$, *and its minimal polynomial be given by* (2.5). *Let $b\in\mathscr{R}(A^k)$. For every $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A^s)$ $(2\leqslant s\leqslant k)$, Krylov iterative sequence $\{x_m\}_{m\geqslant 0}$ converges to $A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ if and only if*

$$\lim_{m\to\infty}\psi_m(A)v=0,\quad \forall v\in\mathscr{R}(A^k);\qquad \lim_{m\to\infty}(I-A^{\mathrm{D}}A)(I-\psi_m(A))v=0,\quad \forall v\in\mathscr{N}(A^s), \tag{3.11}$$

or, equivalently,

$$\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0,\quad 1\leqslant i\leqslant q,\quad 0\leqslant j\leqslant q_i-1;\quad \lim_{m\to\infty}\psi_m^{(j)}(0)=0,\quad 1\leqslant j\leqslant s-1. \tag{3.12}$$

Proof. For every $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A^s)$, we can write it as $x_0=\hat{x}_0+\tilde{x}_0$, $\hat{x}_0\in\mathscr{R}(A^{k-1})$, $\tilde{x}_0\in\mathscr{N}(A^s)$. Let $e_m=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0-x_m$, then

$$e_0=A^{\mathrm{D}}(b-Ax_0)\in\mathscr{R}(A^{\mathrm{D}})=\mathscr{R}(A^k) \tag{3.13}$$

and $r_0=b-A\hat{x}_0-A\tilde{x}_0$, then $(I-A^{\mathrm{D}}A)r_0=-(I-A^{\mathrm{D}}A)A\tilde{x}_0$. Since

$$\begin{aligned}e_m&=\psi_m(A)e_0-\varphi_{m-1}(A)(I-A^{\mathrm{D}}A)r_0\\&=\psi_m(A)e_0+(I-A^{\mathrm{D}}A)A\varphi_{m-1}(A)\tilde{x}_0\\&=\psi_m(A)e_0+(I-A^{\mathrm{D}}A)(I-\psi_m(A))\tilde{x}_0.\end{aligned} \tag{3.14}$$

So, for every $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A^s)$, $\lim_{m\to\infty}x_m=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ is equivalent to $\lim_{m\to\infty}\psi_m(A)v=0$, $\forall v\in\mathscr{R}(A^k)$ and $\lim_{m\to\infty}(I-A^{\mathrm{D}}A)(I-\psi_m(A))v=0$, $\forall v\in\mathscr{N}(A^s)$.

By Theorem 3.2, the first statement of (3.12) is valid. It remains to prove that $\lim_{m\to\infty}(I-A^{\mathrm{D}}A)(I-\psi_m(A))v=0$, $\forall v\in\mathscr{N}(A^s)$, is equivalent to $\lim_{m\to\infty}\psi_m^{(j)}(0)=0$, $1\leqslant j\leqslant s-1$. From (2.3) and (2.4), it is easy to show that $\lim_{m\to\infty}(I-A^{\mathrm{D}}A)(I-\psi_m(A))v=0$, $\forall v\in\mathscr{N}(A^s)$, is equivalent to $\lim_{m\to\infty}N\varphi_{m-1}(N)v=0$, $\forall v\in\mathscr{N}(N^s)$. Using the Taylor expansion of any polynomial $\varphi(\lambda)=\sum_{j=0}^{\infty}[\varphi^{(j)}(0)/j!]\lambda^j$ and $N^j=0$, for $j\geqslant k$, we obtain

$$N\varphi_{m-1}(N)=\sum_{j=0}^{k-2}\frac{\varphi_{m-1}^{(j)}(0)}{j!}N^{j+1}. \tag{3.15}$$

Hence, for every $v\in\mathscr{N}(N^s)$,

$$N\varphi_{m-1}(N)v=\sum_{j=0}^{s-2}\frac{\varphi_{m-1}^{(j)}(0)}{j!}N^{j+1}v. \tag{3.16}$$

It follows that $\lim_{m\to\infty}N\varphi_{m-1}(N)v=0$, $\forall v\in\mathscr{N}(N^s)$, if and only if $\lim_{m\to\infty}\varphi_{m-1}^{(j)}(0)=0$, $0\leqslant j\leqslant s-2$. So we prove (3.12) in view of the identity $\psi_m^{(j)}(0)=-j\varphi_{m-1}^{(j-1)}(0)$. This completes the proof of this theorem. □

On the other hand, if, for some $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, the iterative sequence $\{x_m\}_{m\geqslant 0}$ of (3.1) converges to a solution of the system (1.1), then this solution must be $A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$.

Corollary 3.4. *Let* $A\in\mathscr{C}^{n\times n}$ *with* index $(A)=k$. *Let* $b\in\mathscr{R}(A^k)$. *Assume that the iterations* (3.1) *converge for some* $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$ *and*

$$z=\lim_{m\to\infty}x_m,$$

is a solution of $Ax=b$, *then* $z=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$.

Proof. Suppose z and $A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ are the solutions of the system (1.1), then $z-A^{\mathrm{D}}b-(I-A^{\mathrm{D}}A)x_0\in\mathscr{N}(A)\subseteq\mathscr{N}(A^k)$. Because

$$x_m=(I-A^{\mathrm{D}}A)x_0+A^{\mathrm{D}}Ax_0+\varphi_{m-1}(A)r_0, \tag{3.17}$$

there follows that $x_m\in(I-A^{\mathrm{D}}A)x_0+\mathscr{R}(A^k)$. Then,

$$z=\lim_{m\to\infty}x_m\in(I-A^{\mathrm{D}}A)x_0+\mathscr{R}(A^k). \tag{3.18}$$

Note that $A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0\in(I-A^{\mathrm{D}}A)x_0+\mathscr{R}(A^k)$, we have

$$z-A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0\in\mathscr{R}(A^k). \tag{3.19}$$

Hence, $z-A^{\mathrm{D}}b-(I-A^{\mathrm{D}}A)x_0\in\mathscr{R}(A^k)\bigcap\mathscr{N}(A^k)=\{0\}$, i.e. $z=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$. □

As a special case index $(A)=1$, we immediately obtain the following corollary.

Corollary 3.5. *Suppose* $A\in\mathscr{C}^{n\times n}$ *with* index $(A)=1$ *and* $b\in\mathscr{R}(A)$. *The minimal polynomial of* A *is given by* (2.5). *Then the following statements are equivalent*:

(i) *The Krylov iterative sequence* $\{x_m\}_{m\geqslant0}$ *converges, for all* $x_0\in\mathscr{C}^n$, *to a solution of the system* (1.1);

(ii) $\lim_{m\to\infty}\psi_m(A)v=0$ *for every* $v\in\mathscr{R}(A)$;

(iii) $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0,\ 1\leqslant i\leqslant q,\ 0\leqslant j\leqslant q_i-1$;

(iv) $\lim_{m\to\infty}\varphi_{m-1}(A)v=A^{\#}v,\ \forall v\in\mathscr{R}(A)$;

(v) $\lim_{m\to\infty}\psi_m(A)=I-AA^{\#}$.

Proof. The equivalence among (i)–(iv) are the direct results of Theorem 3.2. Now, we are in position to show that (ii) is equivalent to (v).

(v)⇒(ii) is obvious. Conversely, suppose (ii) holds. For each $v\in\mathscr{R}(A)$, $\lim_{m\to\infty}\psi_m(A)v=0=(I-AA^{\#})v$; for each $v\in\mathscr{N}(A)$, $\psi_m(A)v=v$, and we have $\lim_{m\to\infty}\psi_m(A)v=v=(I-AA^{\#})v$. Thus $\lim_{m\to\infty}\psi_m(A)=I-AA^{\#}$. □

As well known, the minimal Euclidean norm solution x^* of system (1.1) is the unique solution contained in $\mathscr{R}(A^*)$, and so $x^*=P_{\mathscr{R}(A^*)}A^{\#}b$, where $P_{\mathscr{R}(A^*)}$ denotes the orthogonal projection on $\mathscr{R}(A^*)$. Therefore, the Krylov iterative sequence $\{x_m\}_{m\geqslant0}$ converges to the solution x^* if and only if $x_0\in P_{\mathscr{R}(A^*)}A^{\#}b+\mathscr{R}(A)$. Since $A^{\#}=A^+$ if and only if $\mathscr{R}(A)=\mathscr{R}(A^*)$, we have:

Corollary 3.6. *If* A *is range-Hermitian, that is* $\mathscr{R}(A)=\mathscr{R}(A^*)$, *and if the conditions of Corollary* 3.5 *are satisfied, then the Krylov iterative sequence* $\{x_m\}_{m\geqslant0}$ *converges to the minimal Euclidean norm solution* x^* *if and only if* $x_0\in\mathscr{R}(A)$.

4. The case $b\notin\mathscr{R}(A^k)$, $k=\mathrm{index}(A)$

In this section, we shall extend Theorem 3.2 to the case $b\notin\mathscr{R}(A^k)$. First we need a technical lemma.

Lemma 4.1. *Let $A \in \mathscr{C}^{n\times n}$ with* $\operatorname{index}(A)=k$ *and its minimal polynomial is given by* (2.5). *We further fixed an integer t, $1 \leqslant t \leqslant k$. Then the following four statements are equivalent*:

(i) $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$ *and* $\lim_{m\to\infty}\psi_m^{(j)}(0)=0$, $1\leqslant j\leqslant t$;

(ii) $\lim_{m\to\infty}\psi_m(A)v=(I-A^{\mathrm{D}}A)v$, $\forall v\in\mathscr{R}(A^k)+\mathscr{N}(A^{t+1})$;

(iii) $\lim_{m\to\infty}\varphi_m^{(j)}(\lambda_i)=\lim_{m\to\infty}\frac{(-1)^j j!}{(\lambda_i)^{j+1}}$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$ *and* $\lim_{m\to\infty}\varphi_{m-1}^{(j)}(0)=0$, $0\leqslant j\leqslant t-1$;

(iv) $\lim_{m\to\infty}\varphi_{m-1}(A)v=A^{\mathrm{D}}v$, $\forall v\in\mathscr{R}(A^k)+\mathscr{N}(A^t)$.

Proof. (i)⇔(ii): Denote $A_N=P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}A$ and $A_R=(I-P_{\mathscr{N}(A^k),\mathscr{R}(A^k)})A$, then there holds

$$\psi_m(A)=(I-P_{\mathscr{N}(A^k),\mathscr{R}(A^k)})\psi_m(A_R)+P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}\psi_m(A_N). \tag{4.1}$$

For every $v\in\mathscr{R}(A^k)+\mathscr{N}(A^{t+1})$, we decompose $v=v_1+v_2$, $v_1\in\mathscr{R}(A^k)$ and $v_2\in\mathscr{N}(A^{t+1})\subseteq\mathscr{N}(A^k)$, then

$$\psi_m(A)v=\psi_m(A_R)v_1+P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}\psi_m(A_N)v_2. \tag{4.2}$$

It is easy to see that $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$, is equivalent to $\lim_{m\to\infty}\psi_m(A_R)v_1=0$, $\forall v_1\in\mathscr{R}(A^k)$. Observe that $A_N^j v_2=0$, for $j\geqslant t+1$ if $v_2\in\mathscr{N}(A^{t+1})$. It follows from the Taylor expansion of $\psi_m(A_N)$ that

$$P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}\psi_m(A_N)v_2=\left(P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}+\sum_{j=1}^{t}\frac{\psi_m^{(j)}(0)}{j!}A_N^j\right)v_2. \tag{4.3}$$

If $\lim_{m\to\infty}\psi_m^{(j)}(0)=0$, $1\leqslant j\leqslant t$, then there holds from (4.2) and (4.3)

$$\lim_{m\to\infty}\psi_m(A)v=P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}v_2=v_2=(I-A^{\mathrm{D}}A)v. \tag{4.4}$$

Conversely, if (ii) holds, then $\lim_{m\to\infty}\psi_m(A)v=0$, $\forall v\in\mathscr{R}(A^k)$, which leads to $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$; $\lim_{m\to\infty}\psi_m(A)v=v$, $\forall v\in\mathscr{N}(A^{t+1})$, which implies $\lim_{m\to\infty}\psi_m^{(j)}(0)=0$, $1\leqslant j\leqslant t$.

(i)⇔(iii): It follows directly from the identity $\psi_m^{(j)}(\lambda)=-j\varphi_{m-1}^{(j-1)}(\lambda)-\lambda\varphi_{m-1}^{(j)}(\lambda)$.

(i)⇔(iv): For every $v=v_1+v_2$, $v_1\in\mathscr{R}(A^k)$, $v_2\in\mathscr{N}(A^t)$, we have

$$\varphi_{m-1}(A)v=\varphi_{m-1}(A)v_1+\sum_{j=0}^{t-1}\frac{\varphi_{m-1}^{(j)}(0)}{j!}A_N^j v_2. \tag{4.5}$$

Suppose (i) holds, then $\lim_{m\to\infty}\psi_m(A)v=0$, $\forall v\in\mathscr{R}(A^k)$ and $\lim_{m\to\infty}\varphi_{m-1}^{(j)}(0)=0$, $0\leqslant j\leqslant t-1$. Thus we deduce (cf. Theorem 3.2) that

$$\lim_{m\to\infty}\varphi_{m-1}(A)v=A^{\mathrm{D}}v_1=A^{\mathrm{D}}v \tag{4.6}$$

in that $\mathscr{N}(A^t)\subseteq\mathscr{N}(A^k)$. Conversely, suppose that (iv) holds, then $\lim_{m\to\infty}\psi_m(A)v=\lim_{m\to\infty}(I-A\varphi_{m-1}(A))v=0$, $\forall v\in\mathscr{R}(A^k)$ and $\lim_{m\to\infty}\varphi_{m-1}(A)v=\lim_{m\to\infty}\sum_{j=0}^{t-1}[\varphi_{m-1}^{(j)}(0)/j!]A_N^j v=0$, $\forall v\in\mathscr{N}(A^t)$. Therefore, $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$ and $\lim_{m\to\infty}\varphi_{m-1}^{(j)}(0)=0$, $0\leqslant j\leqslant t-1$, so (i) holds. This concludes the proof. □

From Lemma 4.1 and (3.1) we obtain

Theorem 4.2. *Suppose that* $\operatorname{index}(A)=k$ *and its minimal polynomial is given by* (2.5). *Assume that* $b\in\mathscr{R}(A^{k-t})$, $1\leqslant t\leqslant k$, *and let the conditions of Lemma* 4.1 *be fulfilled. Then the Krylov iterative sequence* $\{x_m\}_{m\geqslant 0}$ *converges, for each* $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A^{t+1})$, *to*

$$x=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0=A^{\mathrm{D}}b+P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}x_0. \tag{4.7}$$

Proof. The proof is analogous to that of Theorem 2 in [7], and we omit it here. □

The vector $x=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ is not a solution of system (1.1) even if $b\in\mathscr{R}(A)\setminus\mathscr{R}(A^k)$. Notice that $r(x)=P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}b-A_Nx_0$.

Corollary 4.3. *If the initial guess* x_0 *satisfies* $A_Nx_0=0$, *then the vector* $x=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ *is a solution of the consistent system* $Ax=(I-P_{\mathscr{N}(A^k),\mathscr{R}(A^k)})b$.

Let $P=P_{\mathscr{N}(A^k),\mathscr{R}(A^k)}$ and define the norm $\|.\|_P$ as in [7]:

$$\|z\|_P^2=(Pz,Pz)+((I-P)z,(I-P)z),\ \forall z\in\mathscr{C}^n. \tag{4.8}$$

Thus, if $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)\subseteq\mathscr{R}(A^{k-1})+\mathscr{N}(A^{t+1})$, we have

$$\|r(x)\|_P^2=(Pb,Pb). \tag{4.9}$$

Observe that for arbitrary $z\in\mathscr{R}(A^k)+\mathscr{N}(A)$, there holds for $r(z)=b-Az$

$$\begin{aligned}\|r(z)\|_P^2&=(Pb,Pb)+((I-P)(b-Az),(I-P)(b-Az))\\&\geqslant\|r(x)\|_P^2,\end{aligned} \tag{4.10}$$

which implies that for every $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, the vector $x=A^{\mathrm{D}}b+Px_0$ is a generalized least-squares solution in $\mathscr{R}(A^k)+\mathscr{N}(A)$ of system (1.1) with respect to the norm $\|.\|_P$. Such a generalized least-squares solution has the form $z=A^{\mathrm{D}}b+w$, $\forall w\in\mathscr{N}(A)$. Indeed,

$$z=u+w,\quad u\in\mathscr{R}(A^k),\quad w\in\mathscr{N}(A), \tag{4.11}$$

is a generalized least-squares solution with respect to $\|.\|_P$ if and only if $b-Az\in\mathscr{N}(A^k)$ in view of (4.10), or, equivalently,

$$b-Au\in\mathscr{N}(A^k). \tag{4.12}$$

Write $b=b_1+b_2$, $b_1\in\mathscr{R}(A^k)$, $b_2\in\mathscr{N}(A^k)$, then (4.12) is equivalent to

$$b_1-Au\in\mathscr{N}(A^k). \tag{4.13}$$

Note that $b_1-Au\in\mathscr{R}(A^k)$, it follows that $b_1-Au=0$ and therefore, $u=A^{\mathrm{D}}b_1=A^{\mathrm{D}}b$. Recall that

$$\|z\|_P^2=(A^{\mathrm{D}}b,A^{\mathrm{D}}b)+(w,w)\geqslant\|A^{\mathrm{D}}b\|_P^2, \tag{4.14}$$

and the equality holds if and only if $w=0$, i.e., $z=A^{\mathrm{D}}b$.

Theorem 4.4. *Keep the conditions of Theorem* 4.2. *Then, for every* $x_0\in\mathscr{R}(A^k)+\mathscr{N}(A)$, *the Krylov iterative sequence* $\{x_m\}_{m\geqslant 0}$ *converges to* $x=A^{\mathrm{D}}b+Px_0$, *which is a generalized least-squares solution* *in* $\mathscr{R}(A^k)+\mathscr{N}(A)$ *of the system* (1.1) *with respect to the norm* $\|.\|_P$ *defined by* (4.8). *Furthermore*, x *is the unique minimal* $\|.\|_P$ *norm generalized least-squares solution in* $\mathscr{R}(A^k)+\mathscr{N}(A)$ *of system* (1.1) *with respect to the norm* $\|.\|_P$ *if and only if* $x_0\in\mathscr{R}(A^k)$, (*e.g.*, $x_0=0$).

For the special case $\operatorname{index}(A)=1$, we have

Corollary 4.5. *Let* $\operatorname{index}(A)=1$ *and* $b\notin\mathscr{R}(A)$. *If* $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$ and $\lim_{m\to\infty}\psi_m'(0)=0$, *then the iterations* $\{x_m\}_{m\geqslant 0}$ *converge to a generalized least-squares solution of the system* (1.1) *with respect to the norm* $\|.\|_P$ *for every* x_0. *Moreover*, $x=A^{\#}b+(I-A^{\#}A)x_0$ *is the unique minimal* $\|.\|_P$ *norm generalized least-squares solution with respect to* $\|.\|_P$ *if and only if* $x_0\in\mathscr{R}(A)$.

Corollary 4.6. *Suppose* A *is range-Hermitian and* $\lim_{m\to\infty}\psi_m^{(j)}(\lambda_i)=0$, $1\leqslant i\leqslant q$, $0\leqslant j\leqslant q_i-1$ *and* $\lim_{m\to\infty}\psi_m'(0)=0$. *Then the Krylov iterative sequence* $\{x_m\}_{m\geqslant 0}$ *converges to a least-squares solution of system* (1.1), *and* $x=A^{\#}b+(I-A^{\#}A)x_0$ *is the minimal Euclidean norm least squares solution if and only if* $x_0\in\mathscr{R}(A)$.

We point that the norm $\|.\|_P$ is generally different from the norm $\|.\|_S$ (cf. Lemma 3.1), although they are the same when S is an orthogonal matrix. However, by (4.14), we can easily show that the Drazin inverse solution $A^{\mathrm{D}}b$ is the minimal $\|.\|_P$ norm solution of the system (1.1) if $b\in\mathscr{R}(A^k)$.

For the vector $x=A^{\#}b$, Wei proved in [22] that

Theorem 4.7. *Let* $A\in\mathscr{C}^{n\times n}$ with $\operatorname{index}(A)=1$, $b\notin\mathscr{R}(A)$, *then the vector* $x=A^{\#}b+(I-A^{\#}A)z$, $\forall z\in\mathscr{C}^n$, *is a generalized least-squares solution of system* (1.1) *with respect to the norm* $\|.\|_S$. *Moreover*, $x=A^{\#}b$ *is the unique minimal* $\|.\|_S$ *norm generalized least-squares solution with respect to the norm* $\|.\|_S$.

Proof. See Lemma 2.4 in [23] and Theorem 3.3.5 in [22]. □

5. Convergence bounds for QMR and TFQMR

The QMR algorithm [12] is based on the implementation proposed in [10] of the nonsymmetric Lanczos process with look-ahead. Here, we collect some properties of this look-ahead Lanczos algorithm that will be used in the sequel. For proofs and a detailed description of the look-ahead process, we refer the reader to [10] and references given there.

Starting with $v_1=r_0$, the look-ahead Lanczos algorithm generates basis vectors V_m with (1.4) that are orthogonal to a second Krylov subspace induced by A^* and s_0, here $s_0\in\mathscr{C}^n$ is a second nonzero starting vector that can be chosen arbitrarily. More precisely, each vector v_m is characterized by a biorthogonality condition of the form

$$(w,v_m)=0 \quad \text{for all } w\in\mathscr{K}_l(A^*,s_0) \text{ where } l=l(m). \tag{5.1}$$

In (5.1), we have $l(m) = m - 1$ if v_m is computed by means of a standard Lanczos step without look-ahead. If v_m is generated within a look-ahead step, then the biorthogonality condition (5.1) is relaxed, in the sense that $l(m) = m_l - 1$, where the index m_l $(< m)$ marks the beginning of the look-ahead step. Typically, only a few look-ahead steps occur, and-except for the contrived examples — their size is small, so that usually $m - m_l \leqslant 3$. Finally, a crucial point of the look-ahead Lanczos process is that the basis vectors $v_1, \ldots, v_m$ with (5.1) can be generated by means of short recurrences. These recurrence relations can be written compactly in the matrix form

$$AV_m = V_{m+1}H_m^{(e)}. \tag{5.2}$$

Here $H_m^{(e)}$ is an $(m+1) \times m$ upper Hessenberg matrix that is always full column rank m.

In exact arithmetic, the look-ahead Lanczos process terminates after a finite number of steps, say L, which is called the termination index. There are three possible breakdowns that lead the Lanczos process to terminate, see [12]. Usually, the termination happens because the Lanczos process has detected an A-invariant Krylov subspace $\mathscr{K}_L(A, r_0)$ or an A^*-invariant Krylov subspace $\mathscr{K}_L(A^*, s_0)$. Moreover, the first case occurs if, and only if, $v_{L+1} = 0$. However, in general, it can not be excluded that the Lanczos process fails, in the sense that it encounters a look-ahead step of infinite length. This is called an incurable breakdown. Such a breakdown is very rare and does not present a problem in practice. In the case of the latter, we define the termination index L as the number of the last iteration before the infinite look-ahead step. It was proved (cf. [12, Proposition. 2.2]) that the eigenvalues of the $L \times L$ matrix H_L are also eigenvalues of A in any case, i.e.,

$$\sigma(H_L) \subseteq \sigma(A), \tag{5.3}$$

where the matrix H_L is obtained from the last Lanczos matrix $H_L^{(e)}$ by deleting the last row. Note that the termination index L satisfies

$$L \leqslant \min\{\dim \mathscr{K}_n(A, r_0), \quad \dim \mathscr{K}_n(A^*, s_0)\} \leqslant n. \tag{5.4}$$

Then, the mth QMR iteration x_m is defined by (1.5), where z_m is determined by solving the following least squares problem

$$\|d_{m+1} - \Omega_{m+1}H_m^{(e)}z_m\|_2 = \min_{z \in \mathscr{C}^m} \|d_{m+1} - \Omega_{m+1}H_m^{(e)}z\|_2, \tag{5.5}$$

where the vector d_{m+1} is given by

$$d_{m+1} = [\omega_1, 0, \cdots, 0]^{\mathrm{T}} \in \mathscr{C}^{m+1}, \tag{5.6}$$

and

$$\Omega_{m+1} = \operatorname{diag}(\omega_1, \cdots, \omega_{m+1}), \quad \omega_j > 0, \quad j = 1, \cdots, m+1, \tag{5.7}$$

is a weighting matrix. Usually, one chooses the weights

$$\omega_j = \|v_j\|_2 \text{ for all } j. \tag{5.8}$$

Using (1.5), (5.2), and the fact $v_1 = r_0$, One easily verifies that the residual vector r_m corresponding to x_m satisfies

$$r_m = V_{m+1}\Omega_{m+1}^{-1}(d_{m+1} - \Omega_{m+1}H_m^{(e)}z_m). \tag{5.9}$$

Based on the discussions above, we now present the main result of this section.

Theorem 5.1. *Let $A \in \mathscr{C}^{n\times n}$ with* $\operatorname{index}(A)=k$, *and $b \in \mathscr{R}(A^k)$. In the QMR algorithm, x_0 is selected from $\mathscr{R}(A^{k-1})+\mathscr{N}(A)$. Let L be the termination index of the look-ahead Lanczos algorithm, and the weighting matrix Ω_m be chosen as in* (5.8). *Define*

$$H=\Omega_L H_L \Omega_L^{-1}$$

and let $\lambda_1,\cdots,\lambda_p$ be the eigenvalues of H. Let

$$J=CHC^{-1} \text{ with } J=\operatorname{diag}(J(\lambda_1),\cdots,J(\lambda_p)),$$

be the Jordan canonical form of H. Then,

$$\sigma(H)\subseteq\sigma(A)\setminus\{0\}, \tag{5.10}$$

and for $m=1,2,\ldots,L-1$, the residual vectors of the QMR method satisfy

$$\|r_m\|_2 \leqslant K(C)\sqrt{m+1}\,\varepsilon^{(m)}\|r_0\|_2, \tag{5.11}$$

where $K(C)=\|C\|_2\|C^{-1}\|_2$ and

$$\varepsilon^{(m)}=\min_{\psi\in\mathscr{P}_m,\,\psi(0)=1}\ \max_{1\leqslant j\leqslant p}\|\psi(J(\lambda_j))\|_2. \tag{5.12}$$

Moreover, if H is diagonalizable, then

$$\varepsilon^{(m)}\leqslant\min_{\psi\in\mathscr{P}_m,\,\psi(0)=1}\ \max_{\lambda\in\sigma(A)\setminus\{0\}}|\psi(\lambda)|. \tag{5.13}$$

Finally, in exact arithmetic, the QMR algorithm terminates after $L\leqslant n-k$. If it terminates with $v_{L+1}=0$, then $x_L=A^{\mathrm{D}}b+(I-A^{\mathrm{D}}A)x_0$ is a solution of the system (1.1); furthermore, if $\mathscr{R}(A)=\mathscr{R}(A^*)$ and $x_0\in\mathscr{R}(A)$, then $x_L=A^+b$ is the minimal Euclidean norm solution of system (1.1).

Proof. With $b\in\mathscr{R}(A^k)$ and $x_0\in\mathscr{R}(A^{k-1})+\mathscr{N}(A)$, it follows that $r_0\in\mathscr{R}(A^k)$ and $\mathscr{R}(V_L)\subseteq\mathscr{R}(A^k)$, and thus we have (5.10) instead of (5.3). To obtain estimates (5.11) and (5.12), one then proceeds exactly as in the proof of Theorem 5 in [8].

Since, by assumption, $\operatorname{rank}(A)\leqslant n-k$, and with $\mathscr{K}_n(A,r_0)\subseteq\mathscr{R}(A)$, we have $\dim\mathscr{K}_n(A,r_0)\leqslant n-k$. Together with (5.4), it follows that

$$L\leqslant n-k. \tag{5.14}$$

Finally, assume that $v_{L+1}=0$. Then least-squares problem (5.5) reduces to a linear system with coefficient matrix $\Omega_L H_L$. By (5.10), this linear system is nonsingular. Therefore, the minimum in (5.5) is zero and hence $r_L=0$. The formula of x_L follows directly from Corollary 3.4; moreover, if $\mathscr{R}(A)=\mathscr{R}(A^*)$ and $x_0\in\mathscr{R}(A)$, then $x_L=A^+b$ is the minimal Euclidean norm least-squares solution of (1.1) by Corollary 3.6. □

Notice that an analogous convergence bound for the TFQMR algorithm can be achieved, which is proceeded exactly as in the proof of Theorem 4.2 in [9], so we do not repeat it again.

To conclude this paper, we point that there are other algorithms to compute the Drazin inverse solution $x=A^{\mathrm{D}}b$, for example, an index splitting method is proposed by Wei in [23]. Also, a perturbation bound for the Drazin solution $x = A^{\mathrm{D}}b$ is presented in [25], see also [22, Theorem 3.3.6] and [24].

Acknowledgements

We would like to thank the referees for their useful suggestions. We are also grateful to Dr. J. Xue who helped us improve the presentation of the original paper.

References

[1] V.A. Barker, Numerical solution of sparse singular systems of equations arising from ergodic Markov chain modeling, Comm. Statist. Stochastic Models 5 (1989) 335–381.
[2] A. Ben-Israel, T.N.E. Greville, Generalized Inverse Theory and Applications, Wiley, New York, 1974.
[3] P.N. Brown, H.F. Walker, GMRES on (nearly) singular systems, SIAM J. Matrix Anal. Appl. 18 (1) (1997) 37–51.
[4] S.L. Campbell, C.D. Meyer Jr, Generalized inverse of linear transformations, Pitman, London, 1979.
[5] Z.H. Cao, Polynomial acceleration methods for solving singular systems of linear equations, J. Comput. Math. 9 (4) (1991) 378–387.
[6] J.-J. Climent, M. Neumann, A. Sidi, A semi-iterative method for real spectrum singular linear systems with an arbitrary index, J. Comput. Appl. Math. 87 (1997) 21–38.
[7] M. Eiermann, I. Marek, W. Niethammer, On the solution of singular linear systems of algebraic equations by semiiterative methods, Numer. Math. 53 (1988) 265–283.
[8] R.W. Freund, Quasi-kernel polynomials and convergence results for quasi-minimal residual iterations. in: D. Braess, L.L. Schumaker (Eds.), Numerical Methods of Approximation Theory, Birkauser, Basel, 1992, pp. 77–95.
[9] R.W. Freund, A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems, SIAM J. Sci. Comput. 14 (1993) 470–482.
[10] R.W. Freund, M.H. Gutknecht, N.M. Nachtigal, An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices, SIAM J. Sci. Comput. 14 (1993) 137–158.
[11] R.W. Freund, M. Hochbruck, On the use of two QMR algorithms for solving singular systems and applications in Markov chain modeling, Numer. Linear Algebra Appl. 1 (4) (1994) 403–420.
[12] R.W. Freund, N.M. Nachtigal, QMR: a quasi-minimal residual method for non-Hermitian linear systems, Numer. Math. 60 (1991) 315–339.
[13] M. Hanke, Iterative consistency: a concept for the solution of singular systems of linear equations, SIAM J. Matrix Anal. Appl. 15 (1994) 569–577.
[14] N.J. Higham, P.A. Knight, Finite precision behavior of stationary iteration for solving singular systems, Linear Algebra Appl. 192 (1993) 165–186.
[15] J.C. Meza, W.W. Symes, Deflated Krylov subspace methods for nearly singular linear systems, J. Optim. Theory Appl. 72 (3) (1992) 441–457.
[16] V. Migállon, J. Penadés, D.B. Szyld, Block two-stage methods for singular systems and Markov chains, Numer. Linear Algebra Appl. 3 (5) (1996) 413–426.
[17] M. Neumann, R.J. Plemmons, Convergent non-negative matrices and iterative methods for consistent linear systems, Numer. Math. 31 (1978) 265–279.
[18] B. Philippe, Y. Saad. W.J. Stewart, Numerical methods in Markov Chain modeling, Oper. Res. 40 (1992) 1156–1179.
[19] Y. Saad, Krylov subspace methods for solving large unsymmetric linear systems, Math. Comput. 37 (1981) 105–126.
[20] Y. Saad, M.H. Schultz, GMRES: a generalized minimal residual method for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (1986) 856–869.
[21] A. Sidi, Development of iterative techniques and extrapolation methods for Drazin inverse solution of consistent or inconsistent singular linear systems, Linear Algebra Appl. 167 (1992) 171–203.
[22] Y.M. Wei, Solving singular linear system and generalized inverse, Ph.D. Thesis, Institute of Mathematics, Fudan University, Shanghai, March 1997.

[23] Y.M. Wei, Index splitting for the Drazin inverse and the singular linear system, Applied Math. Comput. 95 (1998) 115–124.
[24] Y.M. Wei, On the perturbation of the group inverse and oblique projector, Applied Math. Comput. 98 (1999) 29–42.
[25] Y.M. Wei, G.R. Wang, Perturbation theory for the Drazin inverse and its applications, Linear Algebra Appl. 258 (1997) 179–186.